National Aeronautics and Space Administration
Contract No. NASw-6

Technical Release No. 34-68

# "SPACE COMMUNICATIONS"

Eberhardt Rechtin

JET PROPULSION LABORATORY
A Research Facility of
National Aeronautics and Space Administration
Operated by
California Institute of Technology
Pasadena, California
May 1, 1960

CR-50,722

This paper presents the results of one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under Contract No. NASw-6, sponsored by the National Aeronautics and Space Administration.

# SPACE COMMUNICATIONS

## LECTURE I: FEASIBILITY

The feasibility of space communications depends on the distance between transmitter and receiver, the amount of interference, the amount of information to be transmitted, the design of the communications system, and the primary energy source.

### A. Line-of-Sight Communications

Space communications almost always involves line-of-sight transmission, transmission in which the transmitter and the receiver have a clear view of each other. Under these conditions relatively simple radio transmission laws apply (Fig. 1). Energy is radiated outward from the transmitter on a spherical area $4\pi d^2$, where d is the distance between the transmitter and the intercepting receiver. By the use of antennas, energy may be beamed or concentrated in a preferred direction; the amount of concentration is called the effective gain G of the antenna. Thus, if a power $P_t$ is radiated from the transmitter using an antenna of gain $G_t$, the power per unit area p at a distance d is given by

$$p = \frac{G_t P_t}{4\pi d^2}$$

The amount of power that is intercepted by the receiving antenna depends on the effective area of the antenna $A_r$. (Both G and A, of course, are functions of the angle from which the antenna is viewed; for simplicity we shall assume that the viewing angle is down the center of the beam, thus giving the greatest G and A.) The ratio of received power to transmitted power is therefore given by

$$\frac{P_r}{P_t} = \frac{G_t A_r}{4\pi d^2}$$

The gain G of an antenna is physically obtained by use of a radiating area A; the relationship between the effective A and effective G is

$$G = \frac{4\pi A}{\lambda^2}$$

in which $\lambda$ is the wavelength of the radiated power. The gain of an antenna evidently depends on its size in wavelengths. The ratio of powers may thus be rewritten in several forms:

$$\frac{P_r}{P_t} = \frac{A_t A_r}{(\lambda d)^2} = \frac{G_t G_r \lambda^2}{(4\pi d)^2} = \frac{G_t A_r}{4\pi d^2} = \frac{G_r A_t}{4\pi d^2} \quad (1)$$

The question of which form of Eq. (1) should be used depends on whether gain or area is the more significant parameter in a particular application. For example, an isotropic antenna is defined as having unity gain, which makes its effective area a function of frequency. A parabola is better characterized by its effective area which approximates its physical area; the gain of a parabola is therefore a function of frequency. Some of the appropriate relationships between effective gain and effective area are given in Table 1.

Table 1. Power gain G and effective area A of several antennas.

| Antenna | Gain | Effective Area |
|---|---|---|
| Isotropic | 1 | $\lambda^2/4\pi$ |
| Infinitesimal dipole | 1.5 | $1.5\lambda^2/4\pi$ |
| Half-wave dipole | 1.64 | $1.65\lambda^2/4\pi$ |
| Optimum horn | $10.0A/\lambda^2$ | 0.81A |
| Parabola or lens | $(6.3 \text{ to } 7.5)A/\lambda^2$ | (0.5 to 0.6)A |
| Broadside array | $4\pi A/\lambda^2$ (maximum) | A (maximum) |

Table 1 refers to antennas in their normal usage; if the various antennas are scaled appreciably larger or smaller than normal, these relationships no longer hold. For example, if a parabola is made increasingly smaller, the transmission frequency remaining constant, the antenna will eventually become so small that it will look like an isotropic radiator. On the other hand, if the area of the parabola is increased beyond normal size, transmission frequency remaining constant, a point will finally be realized where construction inaccuracies will make it impossible to use the complete physical area efficiently. Stated in different terms, if we attempt to make an antenna larger and larger, a point will be reached where it is no longer possible to make the gain of the antenna any greater without enormously refining the construction tolerances of the antenna. It can be shown that the degradation in effective gain (or area) due to surface irregularities, $\sigma$, is given by

$$\frac{A_{eff}}{A_{\sigma=0}} = \frac{G_{eff}}{G_{\sigma=0}} = e^{-16\pi^2 \sigma^2/\lambda^2} \tag{2}$$

The effect of such degradations is shown in Fig. 2. Antennas of this size are said to be gain-limited. We might expect Earth-bound antennas in a space communications system to be of gain-limited design.

If the space vehicle's attitude cannot be controlled or if the vehicle's antenna cannot be pointed accurately enough, the amount of possible vehicle antenna beaming is also limited. For example, if the vehicle tumbles and rolls violently, the antenna must radiate uniformly into space and hence G is close to unity. The early satellites were close to this category; i.e., the vehicle antennas were gain-limited. Within a very short time, however, space vehicles will be considerably different. Stabilization accurate to a few degrees will be achieved, making the use of antenna beamwidths of a few degrees possible. Such narrow beamwidths are produced only by using significant antenna area. However, it is difficult to assemble large, accurate antennas, even in outer space. For some years to come it will not be possible to make space vehicle antennas as large, as accurate, and as controllable as Earth-bound antennas. As a consequence, the vehicle antennas will be limited by sheer size--in other words, area-limited. In the early stages of space exploration, areas of tens of square meters may be achievable; in the more distant future, areas of several thousand square meters may be possible. Even such areas, however, are small compared with those of Earth-bound antennas.

The appropriate forms for Eq. (1) can therefore be written as follows:

| Application | Appropriate Form of the Equation | |
|---|---|---|
| Space-to-Earth (early satellite) | $G_t A_r / 4\pi d^2$ | (1a) |
| Space-to-Earth (immediate future) | $G_r A_t / 4\pi d^2$ | (1b) |
| Space-to-space | $A_t A_r / (\lambda d)^2$ | (1c) |

## B. The Interference Problem

If there were no interference of any kind, it would be possible to hear any transmitter at any distance by incorporating sufficient amplification in the receiver. Unfortunately, interference is always present and strongly limits the usable range of communications. Interference enters the communications system in several places: (1) in the transmitter where it affects the stability and clarity of the signal in a spurious way; (2) in the transmitter-receiver space link where its effectiveness depends on its intensity, direction of arrival, and spectral (frequency) characteristics; (3) in the input circuitry of the receiver where random motion of electrons produces noise dependent on the temperature of the circuitry and the amplifier bandwidth; and (4) signal generating circuitry in the receiver where its effects are similar to transmitter perturbations. Interference entering on the link and interference generated in the receiver input circuitry are the most important of these possible sources.

Every effort is made to avoid external interference sources by locating the sensitive space communications receivers as far as possible from civilization. As for the few interfering signals that might still be present, either the receiver must be designed to handle them individually or the experimenter must be prepared to read through such interference in the process of data reduction.

External interference may be minimized by careful design of antennas (including antennas that can almost cancel point-source interference by pointing an antenna pattern null in the appropriate direction), by choice of frequency, and by appropriate coding of the transmission to make the signal distinctive.

Internal interference is one of the fundamental limitations in any measurement system. It is true throughout the realm of physics that the final sensitivity of instruments is limited by random internal activity. This limitation can be illustrated by the human senses. Even for people with perfect hearing, the smallest sound that can be heard is limited by the sounds of blood coursing through the arteries in the head. The faintest light that can be seen is limited by the faint spots produced by the moving fluids of the eye. Both the ear and the eye are capable of great amplification of very weak signals; yet infinitesimal signals can be neither heard nor seen because of the limitations of internal interference. Exactly the same phenomenon occurs in radio receivers. In a conventional receiver, the principal source of internal interference is noise produced by the random motion of electrons in the input circuitry. Because conventional amplification at radio frequencies is achieved using vacuum tubes whose basic principle is the violent excitation of electrons, the amount of noise activity at the receiver inputs may be 20 times as great as might be expected from the motion of electrons at ambient temperature.

It can be shown that the noise power produced in a bandwidth $\Delta f$ is proportional to the temperature of the circuitry. In equation form,

$$P_n = kT\ \Delta f \text{ watts} \tag{3}$$

where

$P_n$ = noise power, watts
$k$ = Boltzmann's constant = $1.38 \times 10^{-23}$ w-sec/°K
$T$ = absolute temperature, °K
$\Delta f$ = bandwidth, cps

Two ways of minimizing this source of interference are evident: reducing T and reducing $\Delta f$.

Reduction of $P_n$ by reducing $\Delta f$ is limited by the desired rate of information flow. If the bandwidth is extremely small, information rate is very low. In normal, real-time usage, teletype requires approximately 100 cps, voice requires about 3000 cps, and television requires about $3 \times 10^6$ cps. It is therefore difficult to reduce $\Delta f$ without reducing the usefulness of the radio link itself.

The temperature T is thus a critical factor in determining the feasibility of space communications. It has also a most significant relationship to the equivalent temperature of receivers, of the Earth in the vicinity of the ground receiving antenna, of the atmosphere through which the radiation passes, and of the galaxy. The temperature T to be used in Eq. (3) is the weighted sum of all the different source temperatures, where the weighting factors depend upon how the source is "seen" by the receiving antenna. If the source is internal to the receiver or uniformly surrounds the receiving antennas, the weighting factor is unity. If the source occupies only a fraction of the surroundings, its weighting factor depends upon its size and upon its location in the antenna pattern. By definition, the maximum weighting factor for any source is unity. The weighted temperature $\Delta T$ (the change in T produced by an external source) is given by

$$\Delta T = \iint T_s(\phi, \theta)\ G(\phi, \theta)\ \frac{d\phi\ d\theta}{4\pi} \tag{4}$$

where $\phi$ and $\theta$ are angular coordinates and where both the source temperature $T_s$ and the gain G are functions of these coordinates.

Conventional receivers generate a considerable amount of internal noise, noise which has an equivalent temperature of about 2000°K. Parametric-amplifier receivers have an equivalent temperature of about 100°K; masers have an equivalent temperature of about 10°K. Noise external to the receiver has a temperature which is a strong function of frequency (Fig. 3). Not shown in Fig. 3 is one further external source of noise, the black-body radiation of the Earth. This source is relatively constant with frequency. The amount of noise from this source which enters the receiver is a function of the sidelobe level of the receiving antenna (assuming that the main beam is not pointed even partially into the ground). Antennas whose sidelobe levels are low accept much less of this noise than antennas whose sidelobe levels are high; the former antennas are called low-temperature antennas and are characterized by equivalent temperatures of from 10 to 20°K. (A typical well-made parabola and feed has a temperature of about 100°K.) Including the antenna and the maser receiver together, receiver-antenna systems of 30°K temperature are perhaps achievable in the future; whereas systems of just a few years ago were capable of little less than 2000°K.

The total noise power in the link is the sum of the external noise of Fig. 3 and the receiver-antenna system noise. Two cases are illustrated in Fig. 4, the first case using a conventional 2000°K receiving system. The over-all effects of the use of masers are to encourage the use of higher frequencies and to discourage pointing the ground antenna within 10 deg of the horizon.

Atmospheric noise is seen to increase significantly as the antenna is pointed closer to the horizon (more of the atmosphere is traversed), an effect which is much more important for a maser system than for the earlier 2000°K system. In addition, as the antenna is pointed closer to the horizon, the first sidelobes will intercept the black-body radiation from the Earth and will contribute a further increase (50°K anticipated). Consequently, space communications near the horizon will be degraded, relative to communications above a 10-deg elevation angle, by as much as 6 db for maser systems. The consequences are not serious for deep-space probes inasmuch as the total time during which the communications are degraded is comparatively small; the consequences for low-altitude-satellite communications are much more serious because the satellites spend proportionately more time near the horizon.

A further source of external noise is the noise from the Sun, the Moon, and the planets. Of these sources, the Sun is by far the strongest, and the space communicator's only hope is to avoid it whenever possible. Avoiding the Sun with a maser receiver (equivalent temperature of 30°K) generally means pointing the antenna sufficiently far away in angle so that neither the main beam nor the first sidelobes intercept the Sun. If the main beam of the antenna were pointed directly at the Sun and if the beamwidth of the antenna agreed with the angular diameter of the Sun (015 deg), then the temperature would be 300,000°K at 500 Mc, decreasing to 30,000°K at 5000 Mc (see Ref. 10, 13-15, 17). If the Sun were in the first sidelobes of a typical antenna, the respective temperatures would be 1000°K and 100°K. Therefore, whenever the direction of the probe from the Earth is within about 5 deg of the Sun, space communications with masers will be seriously degraded.

The noise temperature of the Moon is between 200 and 300°K; however, the Moon presents a different problem because in order to accomplish communications with a lunar probe (orbital or landing), the antenna perforce must look directly at the Moon. With ground antennas of the gains discussed earlier, the width of the main beam is comparable to the angular diameter of the Moon; consequently, the contribution of the Moon's temperature (about 300°K) to the system temperature depends upon the ratio of the angular diameter of the Moon (0.5 deg) to the beamwidth of the antenna ($\Delta\theta$ deg to 3 db points) (Eq. 4). Assuming that the Moon is in the center of the beam, the temperature increment is

$$\Delta T = \frac{75}{(\Delta\theta)^2} = 2.8 \times 10^{-3} \, G_{max}$$

providing $\Delta\theta$ lies between 0.5 and 20 deg (gains between $10^2$ and $10^5$). For smaller beamwidths (higher gains), the temperature is given by the temperature of the Moon.

Planetary temperatures in the frequencies of interest are all less than 1000°K; more important, the angular diameters of the planets are so small that the average temperature within the main beam is low. The apparent temperatures of Venus and Mars, for example, are the following:

| Planet | Source Temperature | Apparent Temperature |
|---|---|---|
| Venus | 650°K | less than $0.26/(\Delta\theta)^2$ or $10^{-5}$ G |
| Mars | 300°K | less than $0.75 \times 10^{-2}/(\Delta\theta)^2$ or $3 \times 10^{-7}$ G |

where $\Delta\theta$ and G are the antenna beamwidth and maximum gain, respectively.

The range at which communication is feasible depends on the ratio of received power $P_r$ to interference power in the information bandwidth $\Delta f$. A ratio of unity is defined as "threshold reception." A ratio of 10 is noisy but usable. The threshold reception range for space-Earth communication (the weakest link in a Earth-space-Earth system) is given by

$$d = \sqrt{\frac{P_t A_t G_r}{4\pi k T \, \Delta f}} \tag{5}$$

The usable range is approximately one-third of this value. Generally speaking, immediately available electronics limits two-way space communication to within the solar system ($10^{10}$ miles). The nearest star is $10^{14}$ miles away.

C. Unique Signals and A Priori Information

To exploit the received signal power to the fullest extent requires techniques derived from information theory. In their simplest forms some of the laws of information theory seem almost trivial, yet it is surprising how many conventional systems owe their less-than-optimum performance to violation of one or more of the following rules.

First, there is no point in transmitting what is already known. For example, in taking a picture of the other side of the Moon it is a good bet that in general it will look like this side--mostly white with occasional markings, the same diameter, etc.

Second, surprises or unexpected events constitute information; the communications should emphasize the critical features of the surprise. For example, the surprise in the sequence 000000010000000 is the 1 located eight places from the start. The surprise is "eight" which can be transmitted in binary form as 01000. In this particular case the energy necessary to transmit the coded information "eight" is only one-third that necessary to transmit the original sequence.

Third, each signal should be as unique and readily recognizable as possible; i.e., it should have the least resemblance to (or "correlation" with) the interference or alternate signals possible. Hence, the greater the number of possible signals, the longer each one must be.

Fourth, there is no point in the receiver's asking questions for which it already has, or can derive, the answer. Asking for absolute range, velocity, and angle is largely unnecessary if a measurement has just been made. The receiver should only ask about changes. In other words, the receiver should compare the actual reception with the anticipated reception.

These statements can be made considerably more precise in the particular case in which the interference is uniformly distributed in frequency and is gaussian in amplitude distribution. Then the receiver should determine its best estimate of what the signal should be by multiplying the reception (signal and interference) by the estimate, averaging the product, and inspecting the result. Thus, if the signal is $s(t)$, the interference $n(t)$, and the estimate $s^*(t)$, the receiver should inspect the averaged product $e(t)$ for the indications necessary to correct the estimate. The first step is to form $e(t)$ (where the overhead bar indicates the averaged product):

$$e(t) = \overline{s^*(t)\,[s(t) + n(t)]} \tag{6}$$

When $s^*(t)$ exactly equals $s(t)$, the result will be a maximum. By inspecting $e(t)$ it is therefore possible to determine what corrections should be made in $s^*(t)$ to make it a better estimate. The block diagram form is given in Fig. 5. This type of signal-comparing or "correlation" receiver has several alternate forms (including a non-feedback form called the matched filter), but the principle of all the various forms is the same--inspection of an averaged product. The technique has been used with success in improving FM reception, tracking satellites (Microlock), creating interference-resistant systems, and analyzing vibration phenomena. Evidently, the communications system should be so designed that

$$\overline{s(t)\,n(t)} \text{ is a minimum with respect to } \overline{s^2(t)}$$

$$\overline{s_i(t)\,s_j(t)} \text{ is a minimum with respect to } \overline{s^2(t)}$$

$$\text{and } \overline{s(t)\,s(t \pm \Delta t)} \text{ is relatively well behaved}$$

The last requirement insures that $s^*(t)$ can be kept in proper time relationship with $s(t)$, i.e., that the receiver is properly synchronized with the received signal.

We now come to the essential compromise in all communications systems, the compromise that limits the fidelity of voice links, the error-free operation of teletypes, and the guidance accuracy of radio-controlled vehicles. In the correlation receiver just described, if the estimated $s^*(t)$ always exactly agrees with $s(t)$ and hence needs no correction [or inspection of $e(t)$], then quite obviously we have violated some of the fundamental rules of a good transmission link. Why did we bother to transmit if we knew the answer all along? Certainly no information was passed along over the link. (The maximum information is the one-bit piece of knowledge that the signal was actually transmitted.) Any information must therefore be contained in the variations which the transmitter makes in the anticipated s(t). These variations will appear as variations of e(t). On the other hand, how are these variations distinguished from the product $\overline{s^*(t)\, n(t)}$? The answer is that the variations are transmitted so that they result in a measurably different power versus frequency distribution at e(t) than is expected from the interference product. The filtering (selecting) technique that is used to separate the signal variations from the interference is mathematically more complicated than can be covered here (Ref. 19). The general idea, however, is that the filter emphasizes the frequency region where the signal power density exceeds the interference power density and de-emphasizes the frequency region where the reverse is true. In Fig. 6 the optimum filter would emphasize between $f_1$ and $f_2$ and reject elsewhere. To obtain the greatest contrast between the signal variation power density and the interference power density without increasing the signal power, we should make the difference between $f_1$ and $f_2$ as small as possible. But this action either limits the information rate or distorts the signal. Thus, the best compromise between information rate (or distortion) and noise error determines the best system design.

# SPACE COMMUNICATIONS

## LECTURE II: DESIGN CONSIDERATIONS

### A. Power Conversion

The severe weight penalty in space vehicles demands the use of highly efficient transmitters. The coding aspects of this problem have been discussed. In addition, there is the problem of converting primary power, available in such forms as sunlight, chemical bonds, and nuclear energy, into an appropriate signal s(t). Each of these primary power sources has different characteristics and therefore different applications.

The conversion of sunlight into signal energy by solar batteries, heat engines, or the equivalent will supply power (watts) for the lifetime of the equipment. The conversion equipment has weight. With solar batteries now available, a weight factor of 2 watts/kg can be achieved.[1] However, the lifetime of the equipment is limited by micrometeorites and X-rays to 10 years or less. In addition, of course, the available sunlight power per unit area depends on the distance to the Sun. (At $10^8$ km approximately 3.5 kw/$m^2$ is received from the Sun, less than 10% of which is typically converted to electricity.) The power per unit area decreases as the square of the distance from the Sun. Hence, this source of power will decrease as the space vehicle goes in the direction of Mars and will increase in the direction of Venus.

Chemical batteries supply a total amount of energy (watt-hours). Chemical batteries (mercury cells) now available can supply about 100 w-hr/kg, although in this case surge loads are limited. For high surge loads, the weight factor is more nearly 20 w-hr/kg.

[1] H. K. Ziegler, U.S. Army Signal Engineering Laboratories

Atomic batteries supply power using the decay products of radioactive material. Unlike chemical batteries, there is no way of drawing power out "ahead of time"; the process cannot be rushed. In this characteristic the atomic battery resembles the solar battery. On the other hand, the atomic battery has a finite half-life. Atomic battery development is just beginning at this date-- present estimates may change considerably. At present, approximately 1500 w-hr/kg are available; the watt-hours must be used over the half-life of 3-1/2 years. Thus 900 kg of batteries would be required to supply 50 watts for 3 years. A more radical technique using a strontium "hot brick" in a heat engine would have a power capability of 2 kw/kg with a half-life of 25 years. However, handling problems would be severe. A considerable amount of lead shielding might be required.

Another newcomer is the so-called energy cell. Hydrogen and oxygen are stored and then combined at almost any rate to supply power on demand. Like other chemical storage, this cell supplies total energy in watt-hours. The present weight factor is 2000 w-hr/kg.

All these power sources supply power at dc. None supply radio frequency directly. The conversion from dc to radio energy is not notably efficient: 10% is typical; 30% is outstanding. Less-than-perfect efficiency has two consequences: added weight for the power conversion equipment and added weight for heat dissipation. Heat dissipation in a vacuum is not as simple as in the atmosphere where air cooling is available. All excess heat must be radiated. Considerable area may be required to dissipate the heat created by inefficient power conversion.

B. Choice of Transmitting Frequency

The transmitting frequency should be chosen to give the maximum range of communications with a minimum weight required in the vehicle. This maximum range was given by Eq. (5):

$$d = \sqrt{\frac{P_t A_t G_r}{4\pi k T \, \Delta f}}$$

This equation was derived assuming that the vehicle antenna was area-limited and the ground antenna gain-limited. Under these conditions the threshold reception range contains no explicit terms in transmission frequency. Two conclusions may be drawn: (1) a relatively broad range of frequencies may be usable for space communications and (2) the final choice of frequency will probably depend more on the frequency characteristics of the interference and on the state of the art in electronics and space vehicles than on the more obvious terms of Eq. (5).

As might be expected from the first conclusion, the wide variety of studies made on this subject has given answers all the way from 100 Mc to 10,000 Mc. As might be expected from the second conclusion, a detailed result of any study depends on assumptions on the state of the art in electronics and, with considerably less understanding, on assumptions on the state of the art in space vehicles.

At the time that satellites were first planned, practical electronics considerations restricted the choice of frequency to below 300 Mc. The satellites were very low-weight devices, power was at a premium, and the only highly efficient transmitters were transistorized. At that time, transistor transmitters above 200 Mc were not available.

The changes in practical electronics since 1954 are noteworthy. An efficient transmitter at 1000 Mc is being used in the lunar probes; transistor techniques are used in all but the last stage which the state of the art will shortly transistorize as well. Extremely low noise figure receivers of the solid-state maser variety make possible efficient reception to well above 2000 Mc. Solar-energy converters are being developed which will make possible transmitters with 10- to 100-watt capacity at ultrahigh frequencies and microwave frequencies. These new devices are pushing the state of the art at present; by the time interplanetary vehicles are launched, the devices will be nearly commercial. Another development of significance in space communications is the technique of phase-locked receivers for extreme sensitivity and narrow-bandwidth reception at frequencies from 100 Mc to 20, 000 Mc. By following the phase of the incoming signal very accurately, it is possible to cancel out doppler shifts and most oscillator instabilities; the bandwidth necessary to follow the phase is far less than the doppler shifts involved and consequently extremely narrow bandwidth ($\Delta f$) reception is possible. For example, a receiver bandwidth of 20 cps has been used to follow doppler shifts of 10 kc at a carrier frequency of 1000 Mc.

Under the present severe weight and power limitations, it is difficult to generate stable signal power efficiently at frequencies much above 1000 Mc. Within a short time this upper frequency limit may extend to 2000 Mc. However, with the development of heavier vehicles and solar-energy converters, this restriction will be removed, for it will then be possible to move into a new power class (10-100 watts) which permits the use of relatively high efficiency vacuum tubes. The usable frequency range will then extend upward to 10,000 Mc. It should thus be concluded that the state of the art in electronics will not long be a significant limitation to the choice of frequencies for an Earth-space-Earth system.

A more serious factor in the selection of frequency is the dependence of the system temperature $T_s$ on frequency, an effect seen most graphically in Fig. 3. From this figure, frequencies much below 400 Mc would result in system temperatures significantly greater than those achievable using the advanced receiving systems (30°K) aimed more than 10 deg above the horizon and using frequencies above 1000 Mc.

A second factor which favors the frequencies above 1000 Mc is the high cost of large ground antennas. The cost of parabolic antennas is roughly proportional to the 2.7 power of the diameter; therefore, if a 30-m-diameter antenna costs about 1 million dollars, a 130-m antenna will cost about 100 million--a very large figure, indeed. From both Eq. (5) and Fig. 1., it can be seen that the critical parameter for the ground receiving antenna is its gain, not its area. Consequently, a small but accurately surfaced antenna can equal or exceed the performance of a larger but rough-surfaced antenna, provided the best transmitting frequency is selected for each antenna. Because the smaller antenna requires a higher frequency than the larger antenna for best operation, the antenna cost factor strongly favors the higher frequencies.

Based only on the foregoing discussion, the frequency choice would seem to be: the higher the frequency the better. Two factors combine to limit this trend. The first is a sharp increase in atmospheric noise and attenuation above 10,000 Mc due to the presence of water vapor in our atmosphere. The second is the loss of efficiency in using the vehicle antenna area due to surface irregularities, an effect directly comparable to the loss of efficiency in the ground antenna illustrated in Fig. 1. The loss of efficiency occurs when the irregularities are greater than about 5% of a wavelength.

A third factor favors the lower frequencies: the need in certain applications to use omnidirectional antennas on the vehicle. In these applications, the roles of gain and area in the ground and vehicle antennas are reversed (see expressions 1a and 1b for comparison). Under these conditions, the greater the ground antenna area the better (ignoring cost). In order to avoid excessive surface tolerance requirements on the large antennas, this application also means the lower the frequency the better (ignoring external noise). The combination of cost and external noise tends to keep the frequency choice above several hundred megacycles for this particular application.

The gain of the vehicle antenna can also be limited, but less severely, by the ability of the vehicle attitude-control system to aim the directive vehicle antenna to an accuracy better than the angular beamwidth of the antenna. Somewhat by coincidence, however, the state of the art of attitude-control systems is just sufficient to point accurately the vehicle antennas which can be built. Gain-limited vehicle antennas will have gains of about 50 to 60 db and consequently beamwidths of several tenths of a degree. It appears reasonable to expect attitude-control systems capable of this accuracy to be used for space control systems using star references.

For all of the above reasons, it can be concluded that the frequency choice for communications to the Earth from attitude-controlled spacecraft using directive antennas will be greater than 1000 Mc and less than 10, 000 Mc. The resultant ground antenna diameters will be 80 m or less. For communications from vehicles with omnidirectional antennas, the frequencies will be considerably lower with antenna sizes as large as cost and external noise permit.

The choice of frequency for a space-to-space communication can be made in a similar way. It will certainly be more difficult to put extremely sensitive receivers aboard space vehicles than into Earth installations. On the other hand, space-to-space vehicle communications is probably further in the future than Earth-space-Earth communications. The antenna problem for the space-to-space link is alleviated to some extent by the fact that the space vehicle is not continuously rotating as is the Earth. The characteristics of interference in space are reasonably well known for frequencies above approximately 300 Mc; at lower frequencies, the ionosphere surrounding the Earth makes our measurements of interference in space less certain. It will certainly be necessary to exclude as much interference from the galactic plane as possible; therefore, it is desirable to achieve as much gain as possible within the area limitation of the vehicle. These considerations tend to make the frequency choice for space-to-space communications similar to that for space-Earth-space communications. If an area restriction of 1000 $ft^2$ and a gain restriction of $10^5$ are used, the resulting frequency, from Section I-A, is approximately 3000 Mc. Exact choice of frequency is fairly broad for the same reasons as before.

It is certainly true, however, that our lack of knowledge of interference conditions in space leaves the way open for more radical suggestions of frequency choice for space-to-space communications. If, for any reason, interference should be extremely low in certain frequency bands in outer space, it might very well be advantageous to select such frequency bands for space-to-space vehicle communications. As an example, if interference at the lower frequencies were considerably less than what is presently believed, and if the galactic plane concentration of interference were not significant at such low frequencies, it would be possible to use half-wave dipoles at a frequency of several megacycles instead of higher-gain antennas at 1000 Mc to achieve the same ratio of received power to interference power. Such a frequency choice, of course, depends on the interference at this frequency being no worse than at approximately 1000 Mc. At the present time at least this is wishful thinking.

Equally extreme frequency choices can be made in the extremely high frequency region by assuming that internal interference effects can be made extremely small, that receiver bandwidths can be made as small as those at 1000 Mc, that pinpoint beaming of transmission is operationally acceptable, and that extremely high gain antennas can be manufactured in the infrared and near-optical frequency ranges. At the present time few of these assumptions are realistic. However, as has happened many times before, the validity of these assumptions can be markedly changed by the introduction of a new receiver technique (such as the maser) for the discovery of markedly different electromagnetic effects in space than were anticipated (such as the effects so rapidly discovered by the first Earth satellites).

C. Description of a Typical 1958 Space Communications System

In the past several years, a variety of space communications systems have been constructed and successfully used, from the extremely simple 20-Mc Russian system to the most recent systems used aboard the Pioneers, the Tiros, and the Transit. It is worthwhile to describe one of these early systems in order to predict the future more clearly.

The mission of the Pioneer IV was to communicate significant scientific information from the vicinity of the Moon in 1958. The available vehicle was the Juno II (Fig. 7). The payload for this vehicle was very small (Fig. 8). Because of severe weight limitations imposed by the launching vehicle upon the payload, the communications system was limited to an essentially omnidirectional vehicle antenna and less than 200 mw of radiated power. The antenna selected for reception was a modified, 26-m-diameter, radio-astronomy design (Fig. 9). The receivers were conventional, resulting in a system temperature of about 2000°K. Tracking data were teletyped from a variety of sites to the Jet Propulsion Laboratory, where the orbit was computed (Fig. 10). The principal telemetry data from this flight was cosmic-ray information. This information was obtained from sea level to an altitude of 720, 000 km. The most interesting portion of this data was obtained during the first 80, 000 km, a portion of which is shown in Fig. 11.

# SPACE COMMUNICATIONS

## LECTURE III: FUTURE SYSTEMS

### A. Future Satellite and Deep-Space Communications

One of the most straightforward ways to predict the future is to predict how various critical parameters will change and then to demonstrate the effect on reception range for a particular communications system such as Pioneer IV (Table 2). In this particular extrapolation, much of the increased capability was achieved by increasing the vehicle antenna gain. Another alternative would have been to use a higher-gain ground antenna without changing the vehicle antenna gain. This latter solution was used by the Space Technology Laboratory for Pioneer V; in addition, Pioneer V was equipped with a 150-watt transmitter. In order to conserve weight, the transmitter was operated only intermittently.

Carrying the extrapolation further, it is possible to specify a variety of illustrative communications systems (Table 3). Each of these applications has unique problems which influence the numerical values chosen for critical parameters. All systems are calculated using the following equation:

$$\frac{P_r}{N} = \frac{\text{signal power in video bandwidth}}{\text{noise power in video bandwidth}} = \frac{P_t\, A_t\, G_r}{4\pi d^2\, k\, T_s\, (2\Delta f')}$$

where

$P_t$ = radiated power from the vehicle, watts

$A_t$ = effective area of the vehicle antenna, square meters

Table 2. Long-term system capabilities

| Characteristic | Capability | | |
|---|---|---|---|
| | 1958 | 1960 | 1962 |
| Transmitter power, watts | 0.1 | 10 | 100 |
| Vehicle antenna gain, db | 6 | 16 | 36 |
| Receiver sensitivity: | | | |
| noise temperature, °K | 2000 | 400 | 40 |
| bandwidth, cps | 60 | 30 | 30 |
| Earth-vehicle range for 10-db S/N ratio, km | $6 \times 10^5$ | $6 \times 10^7$ | $6 \times 10^9$ |

Table 3. Illustrative communications systems

| Parameter | Weather Satellite | 24-hour Multi-Channel TV Relay | Lunar Orbiter with TV | Lunar Lander with TV | Mars Orbiter with TV | Cosmic-Ray Probe at Edge of Solar System |
|---|---|---|---|---|---|---|
| Range, km | $4 \times 10^3$ | $4 \times 10^4$ | $4 \times 10^5$ | $4 \times 10^5$ | $4 \times 10^8$ | $4 \times 10^{10}$ |
| Ground antenna gain | $10^{3*}$ | $3.5 \times 10^5$ | $4 \times 10^{4*}$ | $10^6$ | $10^6$ | $10^6$ |
| Vehicle antenna area, $m^2$ | 0.05 | 0.1 | 7 | 2.5 | 100 | 100 |
| Vehicle antenna beamwidth, deg | omni* | 18* | 2.2 * | 3.6 * | 0.6 * | 0.6 * |
| System temperature, °K | 400 | 300 | 220 | 400 | 100 | 100 |
| Vehicle radiated power, watts | 200 | 100 | 50 | 10 | 150 | 150 |
| Frequency region, kMc | 0.1-0.4 (0.38) | 1-10 (2.3) | 1-10 (2.3) | 1-10 (2.3) | 1-10 (2.3) | 1-10 (2.3) |
| Video bandwidth for $P_r/N = 10^3$, cps | $4 \times 10^6$ | $20 \times 10^6$ | $10^6$ | $10^6$ | $2.5 \times 10^3$ | too small |
| Video bandwidth for $P_r/N = 10$ | not used | not used | $10^8$ | $10^8$ | $2.5 \times 10^5$ | 25 |
| Time for vehicle to reach destination | 20 min | 2 hr | 3 days | 3 days | 200 days | 30 yr |

*Geometric constraint

NOTE: $\Delta\theta = 46 \dfrac{\lambda}{\sqrt{A}} = \dfrac{164}{\sqrt{G}}$

$G_r$ = gain of the ground receiving antenna

d = distance from the vehicle to the ground receiving antenna, meters

k = Boltzmann's constant = $1.38 \times 10^{-23}$ watts/°K/cps

$T_s$ = system temperature, including the effects of all losses

$\Delta f'$ = video, or post-detection, bandwidth (equal to one-half the transmission bandwidth)

These illustrative designs are by no means the only designs which could be made to work, nor are they the exact designs which may be built, but they are sufficiently good that they can be used for studying the effects of parameter variations.

Weather satellite. The weather satellite travels around the Earth at a comparatively low altitude: 500 to 800 km. The slant range from the receiving antenna to the satellite varies considerably during any one pass with the maximum range occurring at the horizon. The satellite is seldom visible to the receiving station for more than 20 min at a time, and consequently, rapid acquisition of the signal and high-speed data transmission are essential. To simplify the station operation, the ground antenna beamwidth is comparatively broad (about 5 deg) with a resultant antenna gain of only 1000. To simplify vehicle operation, an omnidirectional vehicle antenna is chosen, resulting in a comparatively low frequency for transmission. For a variety of reasons, including the need to look close to the horizon, it will be difficult to use system temperatures much less than 400°K. To transmit high-quality coverage in the short times available would require about 4 Mc of bandwidth at a 30-db signal-to-noise ratio. The resultant radiated power of 200 watts would probably be supplied from rechargeable batteries using the Sun as the prime source of energy.

24-hour multi-channel TV relay. The 24-hour satellite, so-called because its rotational period is the same as that of the Earth, provides a fixed object in the sky which can be efficiently used as a communications relay station. The type of communications to be relayed depends upon economics. In this illustration, three to four television channels could be accomodated. Because the relay is fixed in the sky, it is economic to construct fixed ground antenna reflector surfaces of good accuracy and to provide for slight motions of the antenna feed to accommodate such satellite angular motions as might exist. Antenna sizes of about 25 m in diameter would be reasonable in cost. For economic reasons, the antennas must be located comparatively near civilization; consequently it is not practical to demand system temperatures much below 300°K. (For those fortunate installations where local interference can be minimized and the system temperature reduced as a consequence, one economic tradeoff is the direct one between system temperature and ground antenna area.) The vehicle antenna area is limited by a requirement that the antenna beamwidth illuminate the Earth as seen from the satellite. A shaped antenna beam about 18 deg wide is necessary. The resultant radiated power would be provided by a solar-cell power-supply system.

Lunar orbiter with TV. Real-time television can be sent from the Moon with surprisingly little difficulty. In this system, the vehicle antenna is sufficiently broad in beamwidth to illuminate the whole Earth (hence the vehicle antenna is not required to track a moving ground station) and the ground antenna is sufficiently broad in beamwidth to cover the circling orbiter while aiming the beam at the center of the Moon's face (permitting a very simple antenna drive system needing information only from an almanac). About ten frames per second could be transmitted at an acceptable signal-to-noise ratio. Because the television of the Moon is a sufficiently unique problem, a special modulation system might be provided, using more efficient information coding (FM, PCM, digital telemetry, etc.). Most coding techniques achieve their greatest utility when used with signal-to-noise ratios of about 10 db. Applied to the lunar orbiter, coding bandwidths of 100 Mc could be used if desired.

Lunar lander with TV. The surface exploration of the Moon by instruments can be enormously more effective if an observer on the Earth can watch the operation using a television camera mounted on a lunar automobile. The weight and power limitations on such automobiles, and the difficulties of keeping the automobile antenna pointed accurately at the Earth, make it worthwhile to reduce the lunar vehicle capability and to increase the Earth ground capability by increasing ground antenna gain. (Ground tracking is not as severe a problem for the automobile as for the orbiter.) As has been noted in the previous discussion on lunar temperatures, the increased gain of the ground antenna also results in somewhat increased system temperature.

Mars orbiter with facsimile. Although the minimum distance between the Earth and Mars is only about 80 million km, the communication distance to an orbiter of Mars is considerably greater for two reasons: efficient trajectories do not result in arrival at Mars when Mars is at a minimum distance; and it is desired to keep contact for at least one Martian year. The numerical values of the remaining parameters represent no particularly remarkable state of the art, with the possible exception of the vehicle antenna beamwidth.

In this design, the Mars orbiter must keep its antenna pointed at the distant Earth to within a fraction of a degree, a reasonable but not simple task for a space-vehicle attitude-control system. If this task proves too difficult to perform on early flights, the probable tradeoff would be to reduce the antenna area, to increase the vehicle antenna beamwidth as a consequence, and to accept the reduced video bandwidth as a result. An alternate, but much more expensive, tradeoff would be to use a lower frequency of transmission, and then to require a larger ground antenna (but with the same gain) in order to retain the original system capability.

Cosmic-ray probe at edge of solar system. This communications system is essentially the same as that of the Mars orbiter, except that the increased distance to the probe is compensated by a reduction in information transmission rate and quality. The resulting data are comparable to the cosmic-ray information from *Pioneer IV* obtained at the edge of the solar system instead of at the Moon.

## B. The Tracking Problem

Tracking of space probes is essential in order to evaluate launching performance, to permit vernier corrections to the trajectory, and to point highly directional ground antennas. One of the most difficult tracking phases is the search and acquisition phase necessary at a remote station which must track the probe on the first pass after launch. To relieve this problem, the beamwidth of the ground antenna is made at least 10 deg wide (gain of 270). Countdown and pointing information are also transmitted to the remote station as rapidly as earlier data make it possible. Angular tracking information from such stations is not particularly accurate; however, doppler, ranging, and telemetry from such stations can be excellent.

Determining accurate orbits of satellites and space probes is in some ways easier than determining the trajectory of guided missiles. It is possible to take advantage of the very long tracking intervals and the use of widely spaced stations to achieve accuracies of 10 to 50 times the single measurement accuracy of any one station. For example, the ground tracking stations for the Pioneer IV lunar probe were capable of angular accuracies of 1.0 to 0.1 deg (depending on the station), and yet the orbit was accurately determined to an accuracy of better than 0.01 deg within one day after launch. One of the most valuable additions to tracking systems in recent times has been the addition of computers which can not only smooth the data but which can solve for various station errors and eliminate them. A computer treats unknown (but generally constant) station errors simply as further unknown parameters in addition to the orbital parameters. Inasmuch as computers can often solve for several dozen unknown parameters, the computers not only define the orbit but also evaluate the stations. The calculations depend upon the fact that station errors will create apparent effects in the trajectory which could not occur under Kepler's laws of motion for the vehicle. A simple example is an error in the station clock, an error which would make it appear that the probe had instantaneously changed position when the station first began tracking.

The determination of an accurate orbit for a deep-space probe is critically dependent upon measurements made while the probe is moving rapidly upward through the gravitational field of the Earth. The gravitational gradient field is much more significant in this situation than in the case of guided missiles, because the final orbit of the probe is very sharply influenced by the precise way in which the probe passes through the field. As a consequence, crude measurements made while the probe is leaving the Earth can be far more definitive of the final orbit than very precise measurements made when the probe is far away. The same situation is perhaps even more obvious when the probe reaches the gravitational field of a target planet. The probe's range and relative velocity from the Earth are then strongly influenced by the local gravitational field, and such changes provide a far better indication of the position of the probe relative to the planet than would be the case if no gravitational field were present.

Using angular accuracies of a few tenths of a degree, velocity measurements accurate to 1 meter/sec, and ranging information with a precision of 10 m, it is possible to direct space probes to within 20 km on the surface of the Moon or to within a few thousand kilometers of the planets. Having been directed to this accuracy from the Earth, the probes can easily direct themselves to almost any desired location either on the surface of, or in orbit around, the target.

One of the most surprising features of space tracking, therefore, is that tracking accuracy need not be fantastically precise. Instead, the accuracies are comparable to those demanded of a medium-range missile travelling from one point on the surface of the Earth to another.

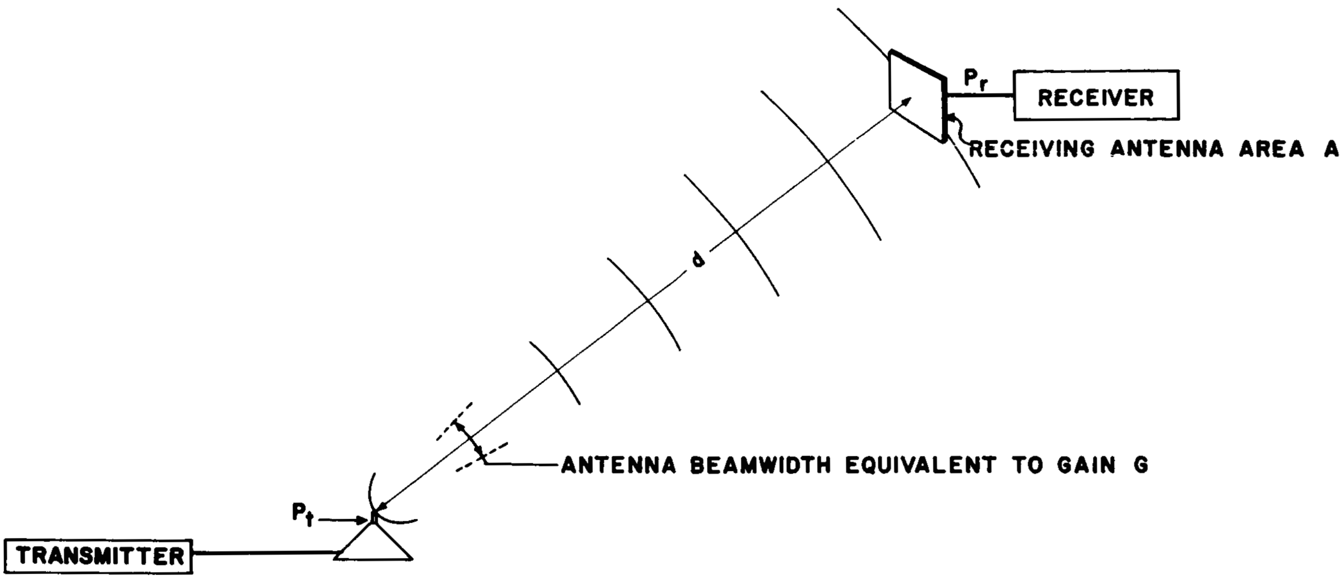
### C. Future Events Having a Major Impact on Communications

Any lecture these days is only up to date for a few years at the most. Here are some events which will make a revision of <u>this</u> lecture necessary.

In the near future, electric propulsion systems will probably come into active use to propel vehicles from Earth satellite orbits out to the planets. It is characteristic of these propulsion systems that they make available a very large source of raw electrical energy--at least kilowatts and, later, megawatts. The space communicator will almost overnight be confronted with more available power than he can possibly convert into radio-frequency energy.

Communications required from vehicles entering a planetary atmosphere at high velocity will be markedly more difficult. The problem will be compounded when the vehicle lands on the surface of the planet and is then carried along by the planet's rotation. Parenthetically, the problem of landing on the Moon is considerably simpler, both because of the lack of atmosphere and because the same surface of the Moon is always presented to the Earth.

In the still more distant future, the advent of vehicles travelling at relativistic velocities will create communications effects which are normally unknown to the space communicator. The doppler shift, of course, will be extremely large. Data transmission rates will be drastically affected. In addition, the apparent direction of the vehicle will be markedly different from its actual direction. As if to confuse the space communicator completely, the apparent antenna pattern of the vehicle antenna will be swept sharply forward. This concentration of energy in the forward direction is a familiar phenomenon to nuclear physicists observing the radiation from a relativistic electron impacting on a plate. The apparent result to an observer on the Earth is a decrease in received power and in signal-to-noise ratio as shown in Fig. 12; but for the space communicator--he will again have entered a new era.




$$\frac{P_r}{P_t} = \frac{G_t A_r}{4 \pi d^2}$$

Fig. 1. Derivation of the range equation

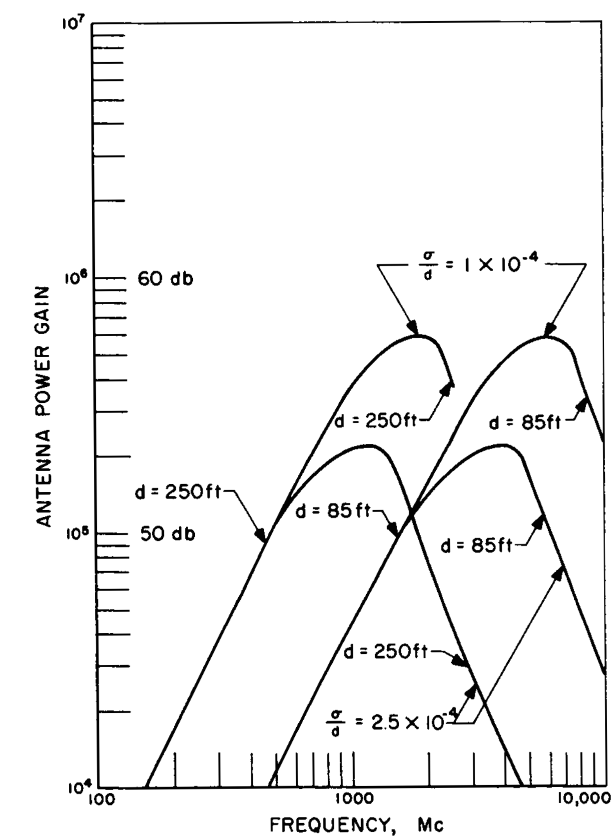



Fig. 2. Antenna gain vs frequency

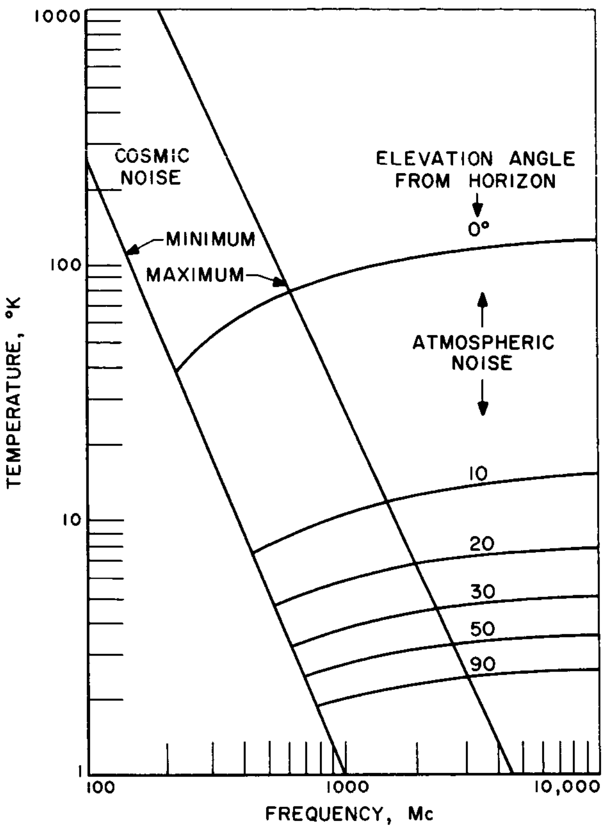



Fig. 3. External noise vs frequency

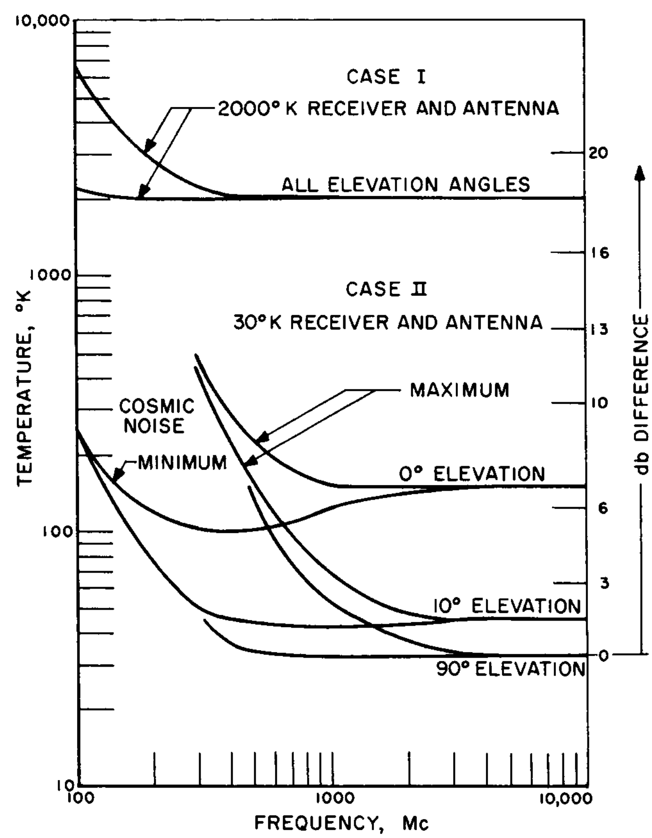
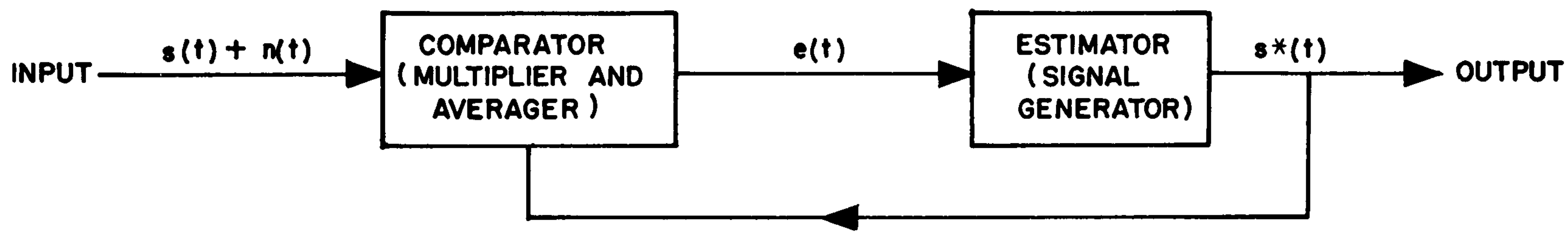



Fig. 4. System temperature vs frequency




DESIGN CRITERIA

$\overline{s(t)\,n(t)}$ IS A MINIMUM WITH RESPECT TO $\overline{s^2(t)}$

$\overline{s_i(t)\,s_j(t)}$ IS A MINIMUM WITH RESPECT TO $\overline{s^2(t)}$

AND $\overline{s(t)\ \ s(t \pm \Delta t)}$ BE RELATIVELY WELL BEHAVED

Fig. 5. Correlation receiver

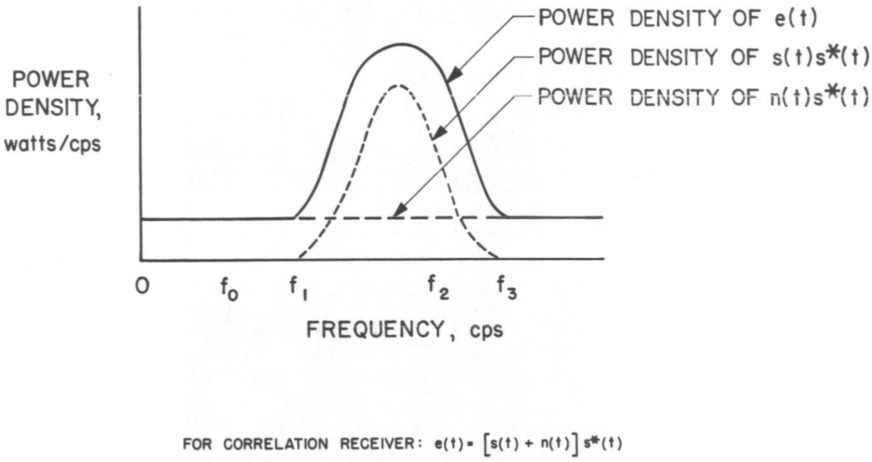



Fig. 6. Power density plot


Fig. 7. Juno II on launcher

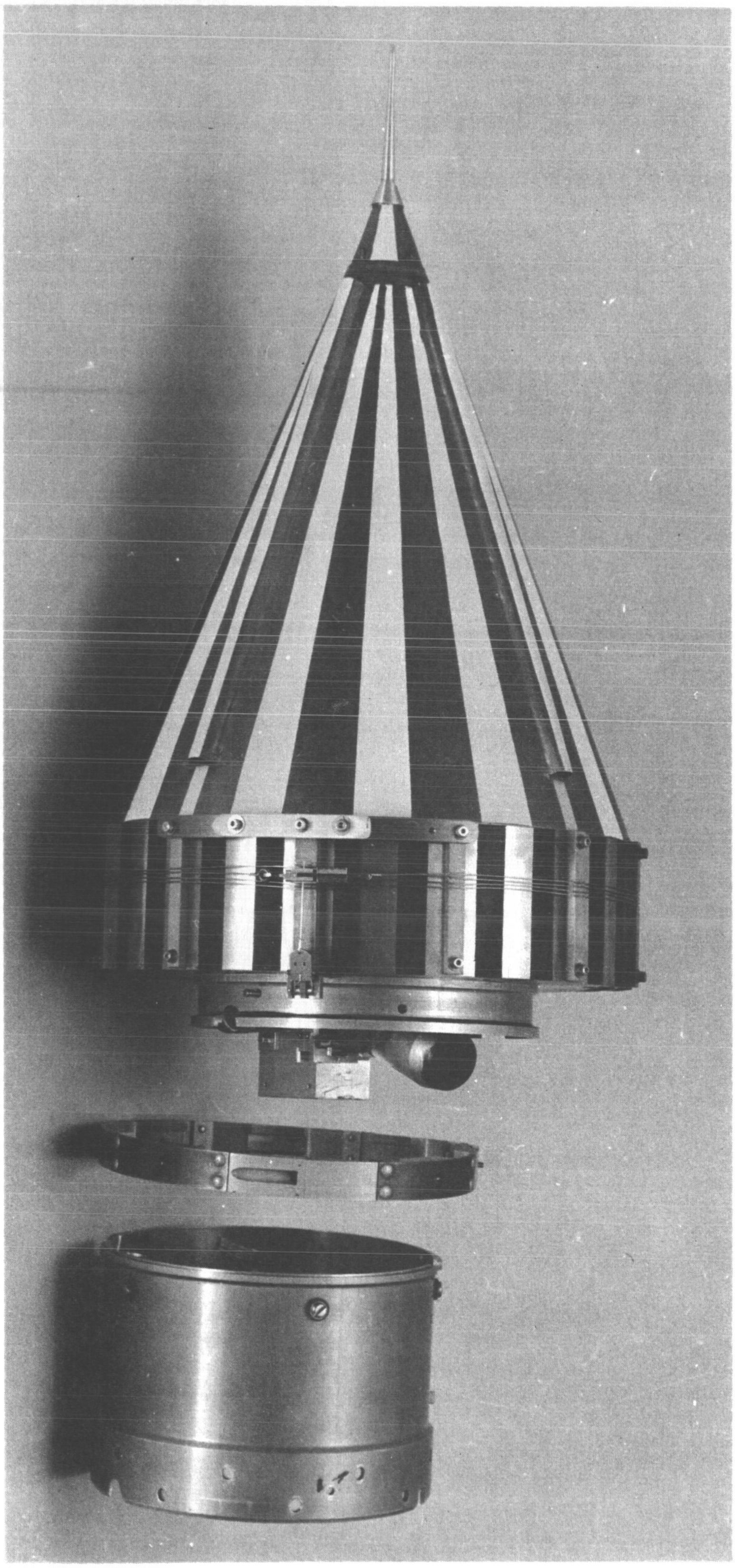


Fig. 8. Pioneer III payload

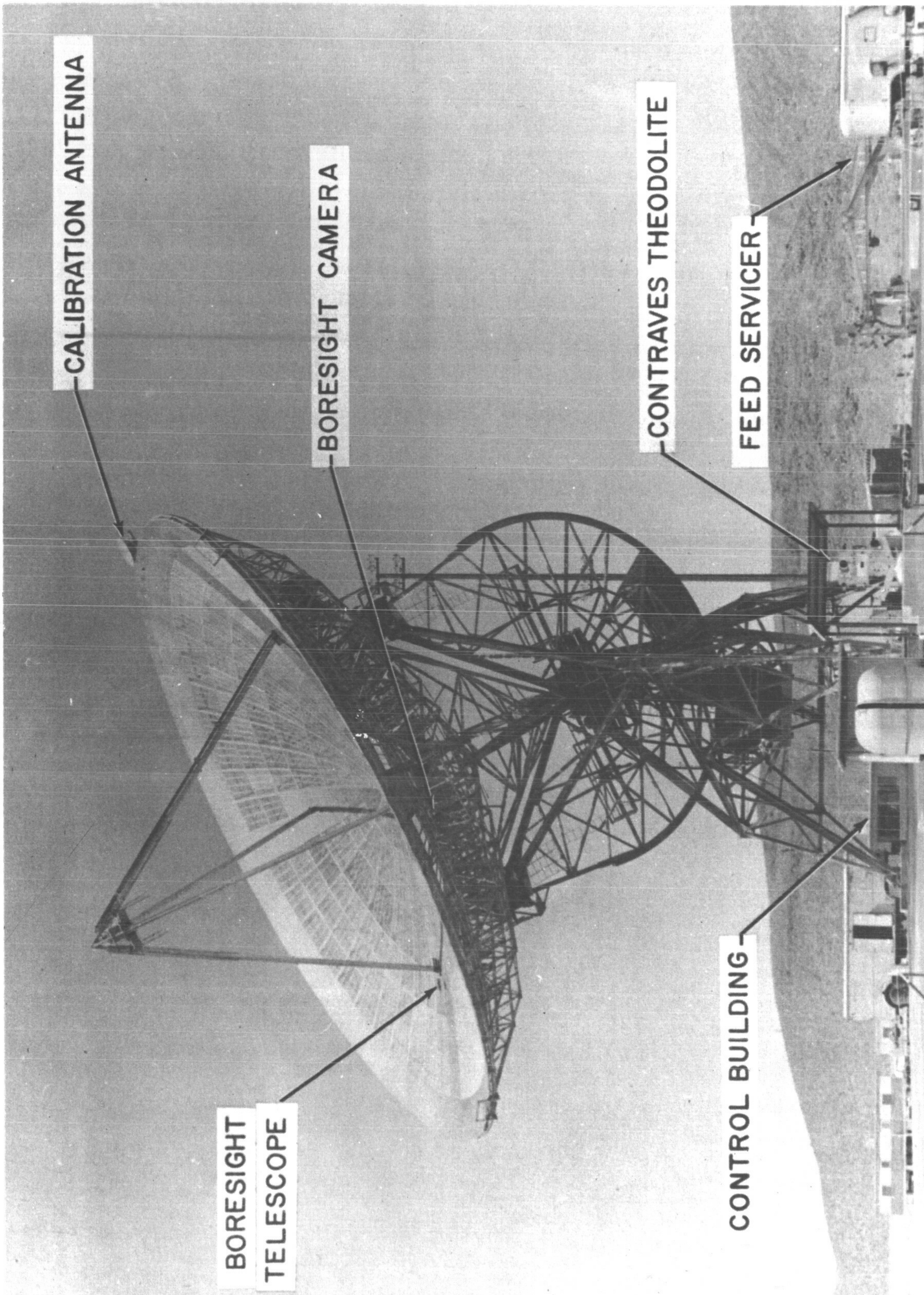



Fig. 9. Goldstone tracking station

| STATION IDENTIFICATION | DATA CONDITION | GMT IN HOURS, MINUTES, SECONDS | GOLDSTONE HOUR ANGLE IN THOUSANDTHS DEGREES | GOLDSTONE DECLINATION ANGLE IN THOUSANDTHS DEGREES | COUNTED DOPPLER FREQUENCY IN CYCLES PER SECOND |
|---|---|---|---|---|---|
| 2 | Ø | 1514Ø1 | 335476 | 336524 | 11337 |
| 2 | Ø | 151411 | 3355Ø8 | 336524 | 11336 |
| 2 | Ø | 151421 | 335556 | 336524 | 11336 |
| 2 | Ø | 151431 | 335648 | 336524 | 11337 |
| 2 | Ø | 151441 | 335736 | 336524 | 11338 |
| 2 | Ø | 151451 | 33582Ø | 336524 | 11336 |
| 2 | Ø | 1515Ø1 | 335848 | 336524 | 11337 |

Fig. 10. Sample of Goldstone data message

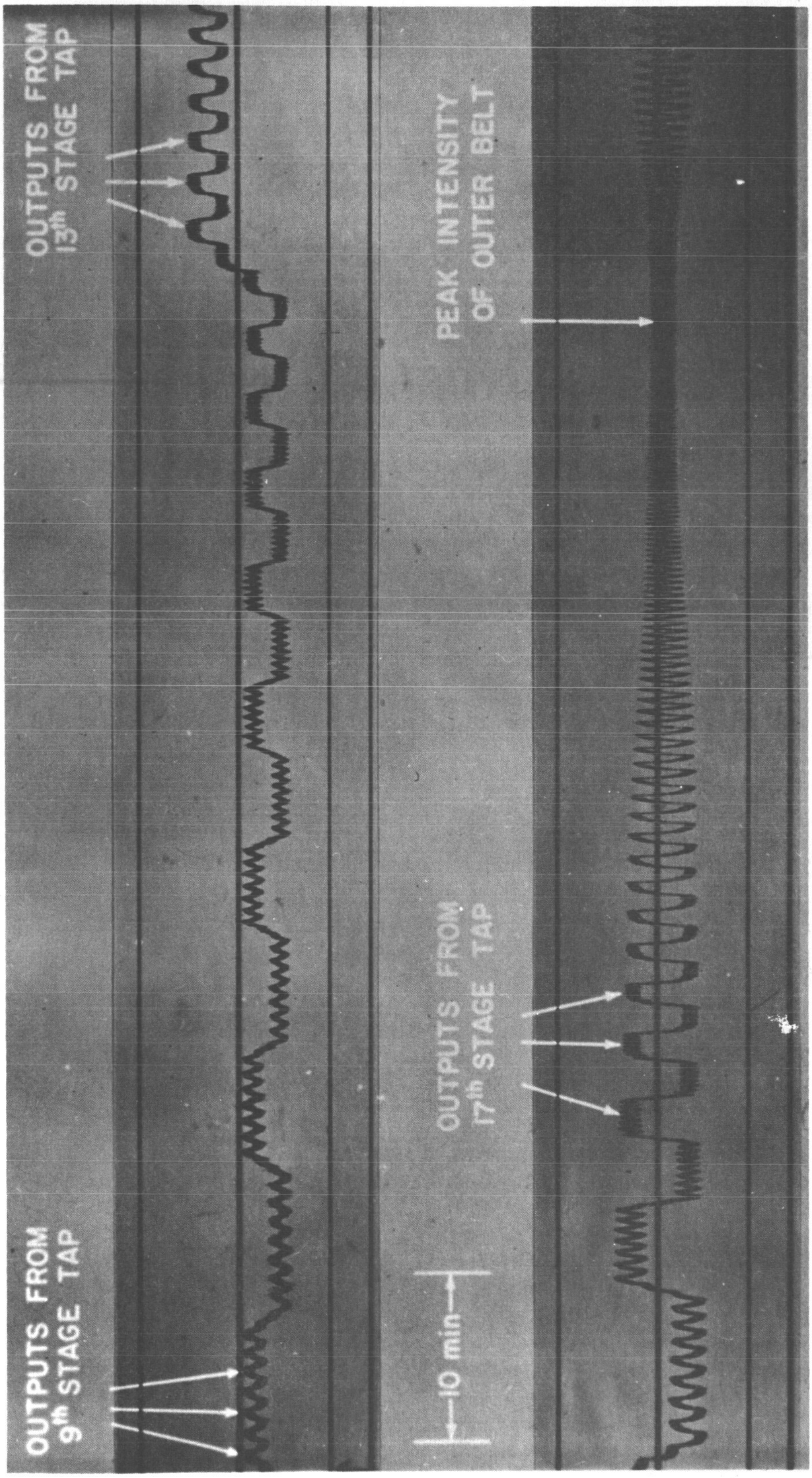



Fig. 11. Cosmic-ray data

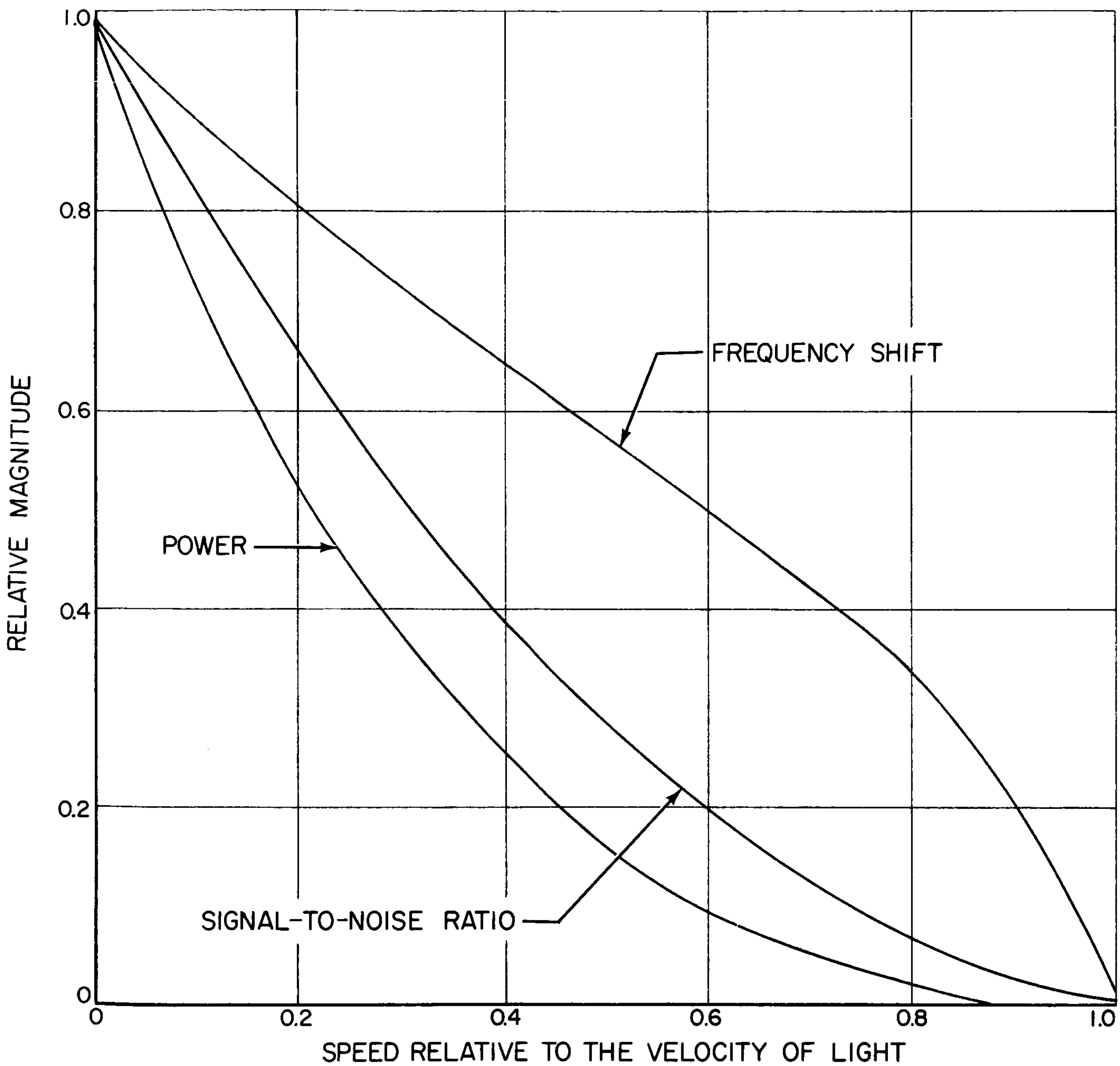



Fig. 12. Relativistic communications

## REFERENCES AND BIBLIOGRAPHY

1. Allen, C. W. "The Quiet and Active Sun," Paper 45 IAU, Symposium No. 4, Radio Astronomy. Cambridge, Cambridge University Press, 1957. Pp. 253-61.

2. Carr, T. D. "Radio Frequency Emission From the Planet Jupiter," Astronomical Journal, Vol. 64 (March 1959), pp. 39-41.

3. Chandra, S. and B. P. Srivastava. "Absorption of Microwaves in Planetary Atmospheres," Zeitschrift für Astrophysik, Vol 47, Part 2 (1959), pp. 127-34.

4. Davenport, W. B., Jr. and W. L. Root. An Introduction to the Theory of Random Signals and Noise, Lexington, Mass., Lincoln Laboratory Publications, 1958.

5. Field, G. B. "The Source of Radiation From Jupiter at Decimeter Wavelengths," Journal of Geophysical Research, Vol. 64, No. 9 (September 1959), pp. 1169-1177.

6. Handbook of Chemistry and Physics, 26th ed., Cleveland, Ohio, Chemical Rubber Publishing Co., 1942.

7. Kraus, J. D. "Planetary and Solar Radio Emission at 11 Meter Wavelength," Proceedings of the IRE, Vol. 46, No. 2 (January 1958), pp. 266-74.

8. Kraus, J. D. and H. C. Ko. Celestial Radio Radiation. Athens, Ohio, State Research Foundation, May 1957.

9. Laning, J. H., Jr. and R. H. Battin. Random Processes in Automatic Control, McGraw-Hill Series, Massachusetts Institute of Technology, 1956.

10. Leighton, Robert B. Principles of Modern Physics, New York, McGraw-Hill, 1958, Chapter 1.

11. Mayer, C. H. "Planetary Radiation at Centimeter Wave Lengths," Astronomical Journal, Vol. 64, No. 2 (March 1959), pp. 43-45.

12. Mayer, C. H., T. P. McCullogh, and R. M. Sloanaker. "Measurements of Planetary Radiation at Centimeter Wavelengths," Proceedings of the IRE, Vol. 46, No. 1 (January 1958), pp. 260-66.

13. Pawsey, J. L. and R. N. Bracewell. Radio Astronomy. Oxford, England, Clarendon Press, 1955.

REFERENCES AND BIBLIOGRAPHY (Cont'd)

14. Pierce, J. R. and R. Kompfner, Proceedings of the IRE, Vol. 47, Part 1 (March 1959), pp. 372-380.

15. Proceedings of the IRE, Vol. 46, Part 1 (January 1958) (complete issue, with special attention to articles beginning on pp. 142, 215, 280, 286, and 293.)

16. Reference Data for Radio Engineers, 4th ed., New York, N. Y., International Telephone and Telegraph Corporation, 1956.

17. Ruze, J. "Effect of Aperture Distribution Errors," Side-Lobe Conference under the auspices of the Radio Divisions I and II, Washington D. C., Naval Research Laboratory, 28 April 1952. (NRL Report No. 4043)

18. Takakura, T. "Synchrotron Radiation and Solar Radio Outbursts at Microwave Frequencies," Paper 101, Paris Symposium on Radio Astronomy, ed. by Ronald N. Bracewell. Stanford, Stanford University Press, 1959. Pp. 562-70.

19. Wiener, N. The Extrapolation, Interpretation, and Smoothing of Stationary Time Series with Engineering Applications, New York, Technology Press and John Wiley and Sons, 1949.